Applied Mathematics Letters 22 (2009) 698–703

Contents lists available at ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

A general framework for the over-relaxed A-proximal point algorithm and applications to inclusion problems

Ram U. Verma

International Publications, 12085 Lake Cypress Circle, Suite I109, Orlando, FL 32828, USA

ARTICLE INFO

Article history:
Received 12 May 2008
Accepted 12 May 2008

Keywords:
Variational inclusions
Maximal monotone mapping
A-maximal monotone mapping
Generalized resolvent operator
Over-relaxed A-proximal point algorithm

ABSTRACT

First, a general framework for the over-relaxed A-proximal point algorithm based on the A-maximal monotonicity is introduced, and second it is applied to the approximation solvability of a general class of nonlinear inclusion problems using the generalized resolvent operator technique. The over-relaxed A-proximal point algorithm is of interest in the sense that it is quite application-oriented, but nontrivial in nature. The results obtained are general in nature.

1. Introduction

Consider a real Hilbert space X with the norm $\|\cdot\|$ and the inner product $\langle\cdot,\cdot\rangle$. Then we consider the general inclusion problem: find a solution to

$$0 \in M(x), \tag{1}$$

where $M : X \to 2^X$ is a set-valued mapping on X.

Recently, the author [1] generalized the over-relaxed proximal point algorithm, investigated by Pennanen [2], and Eckstein and Bertsekas [3], which is based on the celebrated work of Rockafellar [4], to the case of the H-maximal monotonicity [5], while exploring the approximation solvability of (1). Pennanen [2] has shown using the over-relaxed proximal point algorithm and applying a similar approach to Rockafellar [4] by restricting M^{-1} to be locally Lipschitz continuous and by strengthening error tolerance that the sequence converges linearly to a solution of (1).

In this work, we intend to develop a general framework for the over-relaxed proximal point algorithm in the light of the notion of the A-maximal monotonicity of the set-valued map M, which encompasses and unifies all the forms of proximal point algorithms applied in the context of solving general inclusion problems (1) in the literature. The author introduced the notion of A-maximal monotonicity [6], while examining the approximation solvability of the inclusion problems of the form (1) arising from mathematical economics, optimization and control theory, operations research, mathematical finance, mathematical programming, and decision sciences. The A-maximal monotonicity generalizes the existing theory of maximal monotone mappings, including the H-maximal monotonicity [5]. For more literature, we recommend to the reader [1–12].

2. General A-maximal monotonicity

In this section we present some basic properties and auxiliary results on A-maximal monotonicity (also referred to as A-monotonicity in the literature). Let $M : X \to 2^X$ be a multivalued mapping on X. We shall denote both the map M and its

E-mail address: verma99@msn.com.

0893-9659/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aml.2008.05.001

graph by M, that is, the set $\{(x, y) : y \in M(x)\}$. This is equivalent to stating that a mapping is any subset M of $X \times X$, and $M(x) = \{y : (x, y) \in M\}$. If M is single-valued, we shall still use $M(x)$ to represent the unique y such that $(x, y) \in M$ rather than the singleton set $\{y\}$. This interpretation will depend greatly on the context. The domain of a map M is defined (as its projection onto the first argument) by

$$\text{dom}\,(M) = \{x \in X : \exists\, y \in X : (x, y) \in M\} = \{x \in X : M(x) \neq \emptyset\}.$$

The inverse M^{-1} of M is $\{(y, x) : (x, y) \in M\}$.

Definition 2.1. Let $M : X \to 2^X$ be a multivalued mapping on X. The map M is said to be:

(i) (r)-strongly monotone if there exists a positive constant r such that

$$\langle u^* - v^*, u - v\rangle \geq r\|u - v\|^2 \quad \forall (u, u^*), (v, v^*) \in \text{ graph } (M).$$

(ii) (m)-relaxed monotone if there exists a positive constant m such that

$$\langle u^* - v^*, u - v\rangle \geq (-m)\|u - v\|^2 \quad \forall (u, u^*), (v, v^*) \in \text{ graph } (M).$$

(iii) (c)-cocoercive if there exists a positive constant c such that

$$\langle u^* - v^*, u - v\rangle \geq c\|u^* - v^*\|^2 \quad \forall\, (u, u^*), (v, v^*) \in \text{ graph } (M).$$

Definition 2.2 ([6]). Let $A : X \to X$ be a single-valued mapping. The map $M : X \to 2^X$ is said to be A-maximal monotone if:

(i) M is (m)-relaxed monotone for $m > 0$.
(ii) $R(A + \rho M) = X$ for $\rho > 0$.

Definition 2.3 ([6]). Let $A : X \to X$ be an (r)-strongly monotone mapping and let $M : X \to 2^X$ be an A-maximal monotone mapping. Then the generalized resolvent operator $J^M_{\rho,A} : X \to X$ is defined by

$$J^M_{\rho,A}(u) = (A + \rho M)^{-1}(u).$$

Proposition 2.1 ([6]). *Let $A : X \to X$ be an (r)-strongly monotone single-valued mapping and let $M : X \to 2^X$ be an A-maximal monotone mapping. Then $(A + \rho M)$ is maximal monotone for $\rho > 0$.*

3. The over-relaxed *A*-proximal point algorithm

This section deals with an introduction of a generalized version of the over-relaxed proximal point algorithm and its applications to approximation solvability of the inclusion problems of the form (1) based on the A-maximal monotonicity.

Definition 3.1. The map M^{-1}, the inverse of $M : X \to 2^X$, is (c)-Lipschitz continuous at 0 ($c \geq 0$) if there exists a unique solution z^* to $0 \in M(z)$ (equivalently, $M^{-1}(0) = \{z^*\}$) such that

$$\|z - z^*\| \leq \|w\| \quad \text{for } z \in M^{-1}(w) \quad \text{and} \quad \|w\| \leq t(t > 0). \tag{2}$$

Lemma 3.1 ([6]). *Let X be a real Hilbert space, let $A : X \to X$ be (r)-strongly monotone, and let $M : X \to 2^X$ be A-maximal monotone. Then the generalized resolvent operator associated with M and defined by*

$$J^M_{\rho,A}(u) = (A + \rho M)^{-1}(u) \quad \forall\, u \in X,$$

is $(\frac{1}{r-\rho m})$-Lipschitz continuous, where $r - \rho m > 0$.

Proposition 3.1 ([10]). *Let us set $J_k = A - AoJ^M_{\rho,A}oA$. If, in addition,*

$$\langle A(u) - A(v), A(J^M_{\rho,A}(A(u))) - A(J^M_{\rho,A}(A(v)))\rangle \geq \gamma\|A(J^M_{\rho,A}(A(u))) - A(J^M_{\rho,A}(A(v)))\|^2 \quad \text{for } \gamma > \frac{1}{2},$$

then

$$(2\gamma - 1)\|A(J^M_{\rho,A}(A(u))) - A(J^M_{\rho,A}(A(v)))\|^2 + \|J_k(u) - J_k(v)\|^2 \leq \|A(u) - A(v)\|^2 \quad \forall\, u, v \in X. \tag{3}$$

Theorem 3.1. *Let X be a real Hilbert space, let $A : X \to X$ be (r)-strongly monotone, and let $M : X \to 2^X$ be A-maximal monotone. Then the following statements are equivalent:*

(i) *An element $u \in X$ is a solution to (1).*
(ii) *For an $u \in X$, we have*

$$u = J^M_{\rho,A}(A(u)).$$

where

$$J^M_{\rho,A}(u) = (A + \rho M)^{-1}(u).$$

In the following theorem, we apply the generalized over-relaxed A-proximal point algorithm to approximating the solution of (1), and as a result, we establish the linear convergence.

Theorem 3.2. *Let X be a real Hilbert space, let $A : X \to X$ be (r)-strongly monotone and nonexpansive, and let $M : X \to 2^X$ be A-maximal monotone.*

For an arbitrarily chosen initial point x^0, suppose that the sequence $\{x^k\}$ is generated by the over-relaxed proximal point algorithm

$$A(x^{k+1}) = (1 - \alpha_k)A(x^k) + \alpha_k y^k \quad \forall k \geq 0, \tag{4}$$

and y^k satisfies

$$\|y^k - A(J^M_{\rho_k,A}(A(x^k)))\| \leq \delta_k \|y^k - A(x^k)\|,$$

where $J^M_{\rho_k,A} = (A + \rho_k M)^{-1}$, and

$$\{\delta_k\}, \{\alpha_k\}, \{\rho_k\} \subseteq [0, \infty)$$

are scalar sequences. Suppose that the sequence $\{x^k\}$ is bounded and that M^{-1} is (c)-Lipschitz continuous at 0. If, in addition, for $\gamma > \frac{1}{2}$,

$$\langle A(x^k) - A(x^*), A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\rangle \geq \gamma \|A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\|^2, \tag{5}$$

then the sequence $\{x^k\}$ converges linearly to a unique solution x^* of (1) with convergence rate

$$\sqrt{1 - \alpha[2(1 - \gamma d^2) - (1 - (2\gamma - 1)d^2)\alpha]} < 1,$$

where $\alpha = \limsup_{k\to\infty} \alpha_k$, $\rho_k \nearrow \rho$, $d = \limsup_{k\to\infty} d_k = \limsup_{k\to\infty} \sqrt{\frac{c^2}{(2\gamma-1)r^2c^2+\rho_k^2}} < 1$, $\alpha_k^2 + 2\alpha_k(1-\alpha_k)\gamma > 0$, $\alpha_k \geq 1$, $\sum_{k=0}^{\infty} \delta_k < \infty$, and $\delta_k \to 0$.

Proof. Let x^* be a zero of M. We infer from Theorem 3.1 that any solution to (1) is a fixed point of $J^M_{\rho_k,A} oA$. First, in the light of Proposition 3.1, we show

$$\|J^M_{\rho_k,A}(A(x^k)) - x^*\| \leq d_k \|A(x^k) - A(x^*)\|, \tag{6}$$

where $d_k = \sqrt{\frac{c^2}{(2\gamma-1)r^2c^2+\rho_k^2}} < 1$ and $J^M_{\rho_k,A}(A(x^*)) = x^*$. For $J_k = A - AoJ^M_{\rho,A}oA$, and under the assumptions (including (5)), it follows that $A(x^k) - A(J^M_{\rho,A}(A(x^k))) \to 0$. Since $\rho_k^{-1}J_k(x^k) \in M(J^M_{\rho,A}(A(x^k)))$, this implies $J^M_{\rho,A}(A(x^k)) \in M^{-1}(\rho_k^{-1}J_k(x^k))$. Next, applying the Lipschitz condition (2) by setting $w = \rho_k^{-1}J_k(x^k)$ and $z = J_k(x^k) \in M(J^M_{\rho,A}(A(x^k)))$, we have

$$\|J^M_{\rho,A}(A(x^k)) - x^*\| \leq c\|\rho_k^{-1}J_k(x^k)\| \quad \forall k \geq k'. \tag{7}$$

Now applying Proposition 3.1, the (r)-strong monotonicity of A (and hence, A being (r)-expanding) and (7), we get

$$\|J^M_{\rho_k,A}(A(x^k)) - x^*\|^2 \leq d_k^2 \|A(x^k) - A(x^*)\|^2, \tag{8}$$

where $d_k = \sqrt{\frac{c^2}{(2\gamma-1)r^2c^2+\rho_k^2}} < 1$.

Next we start the main part of the proof by using the expression (for all $k \geq 0$)

$$A(z^{k+1}) = (1 - \alpha_k)A(x^k) + \alpha_k A(J^M_{\rho_k,A}(A(x^k))).$$

We begin with estimating (for $\alpha_k \geq 1$) and later using (5), the nonexpansiveness of A, and (8) as follows:

$$\begin{aligned}
\|A(z^{k+1}) - A(x^*)\|^2 &= \|(1-\alpha_k)A(x^k) + \alpha_k A(J^M_{\rho_k,A}(A(x^k))) - [(1-\alpha_k)A(x^*) + \alpha_k A(J^M_{\rho_k,A}(A(x^*)))]\|^2 \\
&= \|(1-\alpha_k)(A(x^k) - A(x^*)) + \alpha_k(A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*))))\|^2 \\
&= (1-\alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k(1-\alpha_k)\langle A(x^k) - A(x^*), A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\rangle \\
&\quad + \alpha_k^2\|A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\|^2 \\
&\leq (1-\alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k(1-\alpha_k)\gamma\|A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\|^2 \\
&\quad + \alpha_k^2\|A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\|^2 \\
&= (1-\alpha_k)^2\|A(x^k) - A(x^*)\|^2 + [\alpha_k^2 + 2\alpha_k(1-\alpha_k)\gamma]\|A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\|^2 \\
&\leq (1-\alpha_k)^2\|A(x^k) - A(x^*)\|^2 + [\alpha_k^2 + 2\alpha_k(1-\alpha_k)\gamma]\|J^M_{\rho_k,A}(A(x^k)) - J^M_{\rho_k,A}(A(x^*))\|^2 \\
&\leq ((1-\alpha_k)^2 + [\alpha_k^2 + 2\alpha_k(1-\alpha_k)\gamma]d_k^2)\|A(x^k) - A(x^*)\|^2.
\end{aligned}$$

where $\alpha_k^2 + 2\alpha_k(1-\alpha_k)\gamma > 0$, $\alpha_k \geq 1$.

Thus, we have

$$\|A(z^{k+1}) - A(x^*)\| \leq \theta_k\|A(x^k) - A(x^*)\|, \tag{9}$$

where

$$\theta_k = \sqrt{1 - \alpha_k[2(1-\gamma d_k^2) - (1-(2\gamma-1)d_k^2)\alpha_k]} < 1,$$

where $\alpha_k^2 + 2\alpha_k(1-\alpha_k)\gamma > 0$, $\alpha_k \geq 1$, $\sum_{k=0}^{\infty}\delta_k < \infty$, $\delta_k \to 0$, and $d_k = \sqrt{\frac{c^2}{(2\gamma-1)r^2c^2+\rho_k^2}} < 1$.

Since $A(x^{k+1}) = (1-\alpha_k)A(x^k) + \alpha_k y^k$, we have $A(x^{k+1}) - A(x^k) = \alpha_k(y^k - A(x^k))$.

It follows that

$$\begin{aligned}
\|A(x^{k+1}) - A(z^{k+1})\| &= \|(1-\alpha_k)A(x^k) + \alpha_k y^k - [(1-\alpha_k)A(x^k) + \alpha_k J^M_{\rho_k,A}(A(x^k))]\| \\
&= \|\alpha_k(y^k - J^M_{\rho,A}(A(x^k)))\| \\
&\leq \alpha_k\delta_k\|y^k - A(x^k)\|.
\end{aligned}$$

Next, we estimate using the above arguments that

$$\begin{aligned}
\|A(x^{k+1}) - A(x^*)\| &\leq \|A(z^{k+1}) - A(x^*)\| + \|A(x^{k+1}) - A(z^{k+1})\| \\
&\leq \|A(z^{k+1}) - A(x^*)\| + \alpha_k\delta_k\|y^k - A(x^k)\| \\
&\leq \|A(z^{k+1}) - A(x^*)\| + \delta_k\|A(x^{k+1}) - A(x^k)\| \\
&\leq \|A(z^{k+1}) - A(x^*)\| + \delta_k\|A(x^{k+1}) - A(x^*)\| + \delta_k\|A(x^k) - A(x^*)\|.
\end{aligned} \tag{10}$$

This implies from (10) on applying (9) that

$$\begin{aligned}
(1-\delta_k)\|A(x^{k+1}) - A(x^*)\| &\leq \|A(z^{k+1}) - A(x^*)\| + \delta_k\|A(x^k) - A(x^*)\| \\
&\leq \theta_k\|A(x^k) - A(x^*)\| + \delta_k\|A(x^k) - A(x^*)\| \\
&= (\theta_k + \delta_k)\|A(x^k) - A(x^*)\|.
\end{aligned} \tag{11}$$

Therefore, we have

$$\|A(x^{k+1}) - A(x^*)\| \leq \frac{\theta_k + \delta_k}{1-\delta_k}\|A(x^k) - A(x^*)\|. \tag{12}$$

Since A is (r)-strongly monotone (and hence, $\|A(x) - A(y)\| \geq r\|x-y\|$), this implies from (12) that the sequence $\{x^k\}$ converges strongly to x^* for

$$\theta_k = \sqrt{1 - \alpha_k[2(1-\gamma d_k^2) - (1-(2\gamma-1)d_k^2)\alpha_k]} < 1,$$

where $\alpha_k^2 + 2\alpha_k(1-\alpha_k)\gamma > 0$, $\alpha_k \geq 1$, $\sum_{k=0}^{\infty}\delta_k < \infty$, $\delta_k \to 0$, and $d_k = \sqrt{\frac{c^2}{(2\gamma-1)r^2c^2+\rho_k^2}} < 1$.

Hence, we use the expression

$$\limsup \frac{\theta_k + \delta_k}{1 - \delta_k} = \limsup \theta_k = \sqrt{1 - \alpha[2(1 - \gamma d^2) - (1 - (2\gamma - 1)d^2)\alpha]} < 1,$$

where $\alpha = \limsup_{k\to\infty} \alpha_k$, $\rho_k \nearrow \rho$, $d = \limsup_{k\to\infty} d_k = \limsup_{k\to\infty} \sqrt{\frac{c^2}{(2\gamma-1)r^2c^2+\rho_k^2}} < 1$. $\square$

For $\gamma = 1$ in Theorem 3.2, we have

Corollary 3.1. *Let X be a real Hilbert space, let $A : X \to X$ be (r)-strongly monotone and nonexpansive, and let $M : X \to 2^X$ be A-maximal monotone. For an arbitrarily chosen initial point x^0, suppose that the sequence $\{x^k\}$ is generated by the over-relaxed proximal point algorithm*

$$A(x^{k+1}) = (1 - \alpha_k)A(x^k) + \alpha_k y^k \quad \forall k \geq 0, \tag{13}$$

and y^k satisfies

$$\|y^k - A(J^M_{\rho_k,A}(A(x^k)))\| \leq \delta_k \|y^k - A(x^k)\|,$$

where $J^M_{\rho_k,A} = (A + \rho_k M)^{-1}$, and

$$\{\delta_k\}, \{\alpha_k\}, \{\rho_k\} \subseteq [0, \infty)$$

are scalar sequences. Suppose that the sequence $\{x^k\}$ is bounded and that M^{-1} is (c)-Lipschitz continuous at 0. If, in addition,

$$\langle A(x^k) - A(x^*), A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\rangle \geq \|A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\|^2, \tag{14}$$

then the sequence $\{x^k\}$ converges linearly to a unique solution x^* of (1) with convergence rate

$$\sqrt{1 - \alpha[(2 - \alpha)(1 - d^2)]} < 1,$$

where $\alpha_k^2 + 2\alpha_k(1 - \alpha_k) > 0$, $\alpha_k \geq 1$, $\sum_{k=0}^{\infty} \delta_k < \infty$, $\delta_k \to 0$, and $\alpha = \limsup_{k\to\infty} \alpha_k$, $\rho_k \nearrow \rho$ and $d = \limsup_{k\to\infty} d_k = \limsup_{k\to\infty} \sqrt{\frac{c^2}{r^2c^2+\rho_k^2}} < 1$.

When $\gamma = 1$, $A = I$ in Theorem 3.2, we have ([2], Proposition 2):

Corollary 3.2. *Let X be a real Hilbert space, and let $M : X \to 2^X$ be maximal monotone. For an arbitrarily chosen initial point x^0, suppose that the sequence $\{x^k\}$ is generated by the over-relaxed proximal point algorithm*

$$x^{k+1} = (1 - \alpha_k)x^k + \alpha_k y^k \quad \forall k \geq 0, \tag{15}$$

and y^k satisfies

$$\|y^k - J^M_{\rho_k}(x^k)\| \leq \delta_k \|y^k - x^k\|,$$

where $J^M_{\rho_k} = (I + \rho_k M)^{-1}$, and

$$\{\delta_k\}, \{\alpha_k\}, \{\rho_k\} \subseteq [0, \infty)$$

are scalar sequences. Suppose that the sequence $\{x^k\}$ is bounded and that M^{-1} is (c)-Lipschitz continuous at 0. Then the sequence $\{x^k\}$ converges linearly to a unique solution x^* of (1) with convergence rate

$$\sqrt{1 - \alpha(2 - \alpha)(1 - d^2)} < 1,$$

where $\alpha_k^2 + 2\alpha_k(1 - \alpha_k) > 0$, $\alpha_k \geq 1$, $\sum_{k=0}^{\infty} \delta_k < \infty$, $\delta_k \to 0$, and $\alpha = \limsup_{k\to\infty} \alpha_k$, $\rho_k \nearrow \rho$ and $d = \limsup_{k\to\infty} d_k = \limsup_{k\to\infty} \sqrt{\frac{c^2}{c^2+\rho_k^2}} < 1$.

Remark 3.1. In Theorem 3.2, if we drop the (c)-Lipschitz continuity of M^{-1} and apply Lemma 3.1 instead, it seems that the strong convergence could be achieved.

References

[1] R.U. Verma, Generalized Rockafellar's theorem on linear convergence, Nonlinear Functional Analysis and Applications (in press).
[2] T. Pennanen, Local convergence of the proximal point algorithm and multiplier methods without monotonicity, Mathematics of Operations Research 27 (1) (2002) 170–191.
[3] J. Eckstein, Bertsekas, on the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators, Mathematical Programming 55 (1992) 293–318.
[4] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM Journal of Control and Optimization 14 (1976) 877–898.
[5] Y.P. Fang, N.J. Huang, H-monotone operators and system of variational inclusions, Communications on Applied Nonlinear Analysis 11 (1) (2004) 93–101.
[6] R.U. Verma, A-monotonicity and its role in nonlinear variational inclusions, Journal of Optimization Theory and Applications 129 (3) (2006) 457–467.
[7] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, Mathematics of Operations Research 1 (1976) 97–116.
[8] H.Y. Lan, J.H. Kim, Y.J. Cho, On a new class of nonlinear A-monotone multivalued variational inclusions, Journal of Mathematical Analysis and Applications 327 (1) (2007) 481–493.
[9] R.U. Verma, Approximation solvability of a class of nonlinear set-valued inclusions involving (A, η)-monotone mappings, Journal of Mathematical Analysis and Applications 337 (2008) 969–975.
[10] R.U. Verma, General inexact A-proximal point algorithm and applications to variational inclusion problems, preprint, 2008.
[11] E. Zeidler, Nonlinear Functional Analysis and its Applications I, Springer-Verlag, New York, 1986.
[12] E. Zeidler, Nonlinear Functional Analysis and its Applications II/B, Springer-Verlag, New York, 1990.